Attachment I

Evaluation of the Hail Detection Algorithm for high-elevation WSR-88D sites

Arthur Witt
NOAA/ERL/National Severe Storms Laboratory

Report prepared for the WSR-88D Operational Support Facility

March 1998

1. Introduction

With the operational implementation of the Build 9 RPG code, the Hail Detection Algorithm (HDA) has undergone real-time evaluation at NWS offices in various regions and/or seasons for which no prior quantitative testing was done. This real-time evaluation has led to reports that the HDA is greatly overforecasting the likelihood of severe hail for those WSR-88D sites located at high elevations. Of the three outputs generated by the HDA — probability of hail (POH), probability of severe hail (POSH), and maximum expected hail size (MEHS) — only the POSH parameter is affected by the radar elevation. This effect occurs via the Warning Threshold Selection Model (WTSM), which currently uses the height above radar level (ARL) of the environmental melting level to determine an optimum "warning threshold" (WT), which is then used in the calculation of POSH. The data used to develop and test the HDA prior to its implementation in Build 9 came almost entirely from low-elevation WSR-88D sites, where there is little difference in the POSH values when the height of the melting level is determined relative to the radar elevation or mean sea level (MSL). However, the impact for WSR-88D sites located at higher elevations (> 1 km MSL) is increasingly large as the site elevation increases. Hence, it appears likely that the overforecasting bias that has been observed at high elevation sites this past year is due to WT values that are too small. A possible solution to this problem is to change the WTSM so that the melting level relative to MSL is used (instead of ARL). To determine if this change would improve HDA performance, WSR-88D data for several storm days from high elevation sites were analyzed.

2. Analysis results

Level II data for eight storm days were analyzed (Table 1). The analysis procedures were the same as those used for previous evaluation of the HDA (Witt et al. 1998). Test results were generated using two reference elevations for the WTSM: ARL and MSL. The combined results for the four storm days from Colorado are shown in Fig. 1. For the current reference elevation of ARL, a large overforecasting bias is apparent. However, by changing the reference elevation to MSL, much of this bias is eliminated.

The combined results for the two storm days from Montana are shown in Fig. 2. Using ARL as the reference elevation, a very large overforecasting bias is apparent. Changing the reference elevation to MSL produces generally minor improvements, except for POSH values of 60% and 70%. However, because of the sparse population across the radar domain (for ranges out to 230 km), the validity of these results is highly questionable, due to deficiencies in ground-truth verification. Because of similar concerns for the data from Tucson, no attempt was made to generate a reliability diagram for that site. Instead, for both the Arizona and Montana data, the POSH values within a 20 minute time window for each severe hail report were analyzed to determine the impact of changing the reference elevation from ARL to MSL (Table 2).

The first Montana case (2 July 1996) was characterized by fairly strong storms, with hail up to baseball size reported. The cell-based VIL values corresponding to each severe hail report are quite high, with values around 50. The ARL-based POSH values are all at or near 100%. The MSL-based POSH values are somewhat lower, but remain $ 50%. In contrast, the second Montana case (1 August 1996) was characterized by relatively weak storms, with only marginally-severe-sized hail reported. The observed cell-based VIL values were quite low on this day, particularly for severe hailstorms occurring during the month of August. Note the first hail report, with a corresponding maximum VIL of only 14. But, even for this marginal event, ARL-based POSH values are still generally quite high (> 60% for three of the four reports). The MSL-based POSH values, on the other hand, are substantially lower (average probabilities < 20%). For the first report, the MSL-based POSH is 0%.

Similar results occur for the two Arizona storm days. For the two severe hail reports associated with strong storms (VIL > 50), both the ARL- and MSL-based POSH values are high, with the MSL-based POSH values - 20 percentage points lower than the ARL-based POSH values. The other two severe hail reports are produced by storm cells with moderately high ARL-based POSH values, which decrease by around a factor of 2 for the MSL-based values.

3. Discussion

The analysis results presented above clearly show a large overforecasting bias for the POSH parameter, as it is currently calculated in Build 9. Now, some of this overforecasting bias is undoubtedly related to verification problems, but some is also likely due to inappropriate parameter values used in the calculation of POSH. Specifically, the equation used to generate the warning threshold, which is a function of the melting level above some reference height, is producing values that are likely too small. Currently, this reference height is the elevation of the RDA. One way to increase the WT values is to use a reference height closer to mean sea level. From the results shown in Figs. 1 and 2, this should improve the reliability of the POSH parameter. However, users of HDA output need to be aware that such a change would likely result in some (and perhaps many) marginal severe hail events being associated with fairly low POSH values (Table 2). But, in order to maintain good reliability (i.e., well-calibrated probability forecasts), some severe hail observations must correspond to the lower POSH values. Preferably, these would be marginal events, which turns out to be the case for the data analyzed here (Table 2). Therefore, as a near-term “quick fix,” based on these test results, it appears appropriate to change the WTSM so that the height of the melting level is determined relative to MSL instead of ARL. This change also has the benefit of normalizing POSH output from multiple WSR-88D sites which happen to scan the same storm at the same time [currently, substantial variations in POSH can occur solely due to differences in radar site elevation (Maddox et al. 1998)].

Although changing the reference elevation to MSL should lead to overall improvements in HDA performance, it is possible that, for some high elevation sites, this adjustment may lead to an actual underforecasting bias. Because of the small number of days analyzed in this study, and the greater verification limitations for the non-Colorado cases, these results should be considered tentative, and most applicable to Colorado. It would appear from this study that the impact on POSH of seasonal variations in the height of the melting level for a radar close to MSL is not the same as variations in the height of the melting level due solely to terrain heights (above MSL). But, there should be some impact of terrain height on the occurrence and size of hail at the surface (due to melting effects). It appears, however, that these factors are much smaller than the seasonal effects, or are masked by verification biases in the data analyzed. Clearly, much additional work is still needed to optimize the HDA for high-elevation (and other) locations. An accurate measure of algorithm performance for a large number of storm days, similar to that done for Florida by Wyatt and Witt (1997), would be highly desirable. To reduce the impact of verification biases and deficiencies, future adaptable parameter sensitivity studies should focus on high-population areas located at different terrain elevations.

Additional avenues for potential enhancement of the HDA include adding information on relevant velocity signatures (e.g., midaltitude rotation and storm-top divergence), broadening the use of reflectivity data (e.g., 3D Severe Hail Index, Bounded Weak Echo Regions), incorporation of more environmental data (e.g., vertical profiles of temperature, humidity and wind), and adding a terrain model. Current plans at NSSL include:

1) Investigating the relationship between midaltitude rotation in a storm and observed hail size. Cloud modeling studies of hail growth have shown that midaltitude rotation plays an important role in the production of very large hail (Miller et al. 1988). Thus, adding this capability to the HDA should improve the accuracy of the maximum expected hail size estimates, along with reducing instances when very large hail (e.g., > 2 inches) is predicted for nonsupercell storms.
2) Calculation of a volumetric (3D) Severe Hail Index (SHI), versus the 1D SHI now used. This should allow for better discrimination between large severe hailstorms and small cells that may be producing little, if any, severe hail, despite having nearly identical 1D vertical reflectivity profiles.
3) Adding a terrain model to the HDA. This should allow us to incorporate varying terrain heights into any new "melting functions" that are developed for the HDA.

Finally, additional testing (using Level II data from Florida and Texas) done within the past year by Wyatt and Witt (1997) has determined that, in summertime environments characterized by relatively high melting levels and low vertical wind shear, there is an overforecasting bias of ∼20 percentage points in the POSH values. A similar bias was also found for data from Ohio and Pennsylvania (Barjenbruch and LaPlante 1997). This bias can be eliminated via a change to one of the adaptable parameters in the POSH equation (Wyatt and Witt 1997). Whether this additional adjustment should be made for high-elevation sites in unclear, since the test results came from sites with elevations near MSL. From Fig. 1, there is still a slight overforecasting bias, suggesting that some additional reduction in the POSH values may be useful. However, because the test results for high-elevation sites involve so few storm days (for any given location), users of HDA output should carefully monitor the algorithm's performance if both of the above mentioned changes are implemented, and reduce the amount of (downward) adjustment to the POSH parameter if a distinct underforecasting bias is observed.

4. References

Barjenbruch, K. M., and R. E. LaPlante, 1997: An assessment of the WSR-88D Hail Detection Algorithm's performance over northern Ohio and northwest Pennsylvania. [Available on-line at http://www.csuohio.edu/nws/hail.htm]

Maddox, R. A., D. R. Bright, W. J. Meyer, and K. W. Howard, 1998: Evaluation of the WSR-88D Hail Algorithm over southeast Arizona. Preprints, *16th Conf. on Wea. Analysis and Forecasting.,* Phoenix, AZ, Amer. Meteor. Soc., 227-232.

Miller, L. J., J. D. Tuttle, and C. A. Knight, 1988: Airflow and hail growth in a severe northern High Plains supercell. *J. Atmos. Sci.,* **45,** 736-762.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, in press.

Wyatt, A., and A. Witt, 1997: The effect of population density on ground-truth verification of the new WSR-88D Hail Detection Algorithm. [Available on-line at http://www.nssl.noaa.gov/swat/NWA97_HDA/HDAstudy.html]

Figure captions

Fig. 1. Reliability diagram for the probability of severe hail (POSH) for the four storm days from Colorado. Numerals adjacent to the plotted points indicate the number of forecasts for that POSH value. The diagonal line represents perfect reliability.

Fig. 2. Same as Fig. 1, except for the two storm days from Montana.

Appendix

Adjusting the Warning Threshold Selection Model (WTSM)

In the Build 9 HDA, the WTSM is defined as

$$WT = aH_0 + ß, \tag{1}$$

where H_0 is the height ARL of the melting level, and a and ß are adaptable parameters. To change the WTSM so that the height of the melting level relative to MSL (instead of ARL) is used, one need only make a simple addition to Equation (1):

$$WT = a(H_0 + RE) + ß, \tag{2}$$

where RE is the radar site elevation. Equation (2) can be rewritten as

$$WT = aH_0 + ?, \tag{3}$$

where ? = aRE + ß. Therefore, replacing the value of ß in Equation (1) with the value ? in Equation (3) will result in the desired adjustment to the WTSM.

Table 1. List of the storm cases analyzed. RS is the radar site, EL is the site elevation (in km MSL), BT and ET are the beginning and ending times of data analysis (in UTC), H_0 is the melting level (in km ARL), NR is the number of hail reports used in the analysis, MS is the maximum reported hail size (in mm), NVS is the number of volume scans analyzed, NAP is the total number of algorithm predictions, and MZ is the maximum reflectivity (in dBZ) for all the storm cells analyzed. The date corresponds to the beginning time. RS locations: EMX is Tucson, AZ; FTG is Denver, CO; MSX is Missoula, MT; and PUX is Pueblo, CO.

RS	EL	DATE	BT	ET	H_0	NR	MS	NVS	NAP	MZ
FTG	1.71	5/12/95	1834	0246	1.7	3	22	68	429	65
FTG		6/22/95	2200	0309	2.7	12	44	48	437	68
FTG		5/22/96	2302	0823	2.75	16	114	99	179	72
PUX	1.62	8/02/96	0322	0438	3.4	3	76	14	90	65
MSX	2.45	7/02/96	2032	0537	2.15	4	70	94	1015	68
MSX		8/01/96	1848	0458	1.85	4	25	101	786	62
EMX	1.62	7/25/96	1909	0257	3.4	2	44	–	–	–
EMX		8/14/96	1934	0323	3.3	2	25	–	–	–

Table 2. Algorithm output associated with each of the hail reports from the Montana and Arizona cases. The POSH (in percent) and VIL (in kg m^{-2}) data are average and maximum values for a 20 minute time window.

Report size (inches)	POSH-ARL (avg/max)	POSH-MSL (avg/max)	Cell-based VIL (avg/max)
	MSX 7/2/96		
1.00	98/100	50/60	43/49
1.00	100/100	63/70	50/53
2.75	100/100	70/70	49/52
1.75	100/100	65/70	48/54
	MSX 8/1/96		
0.75	30/40	0/0	13/14
0.75	64/80	14/30	21/26
1.00	72/90	18/40	27/40
0.75	70/80	18/30	21/28
	EMX 7/25/96		
1.00	34/70	16/40	33/47
1.75	81/100	61/80	63/77
	EMX 8/14/96		
0.75	53/60	28/30	24/44
1.00	70/80	48/60	57/64

Figure 1

Four days from Colorado

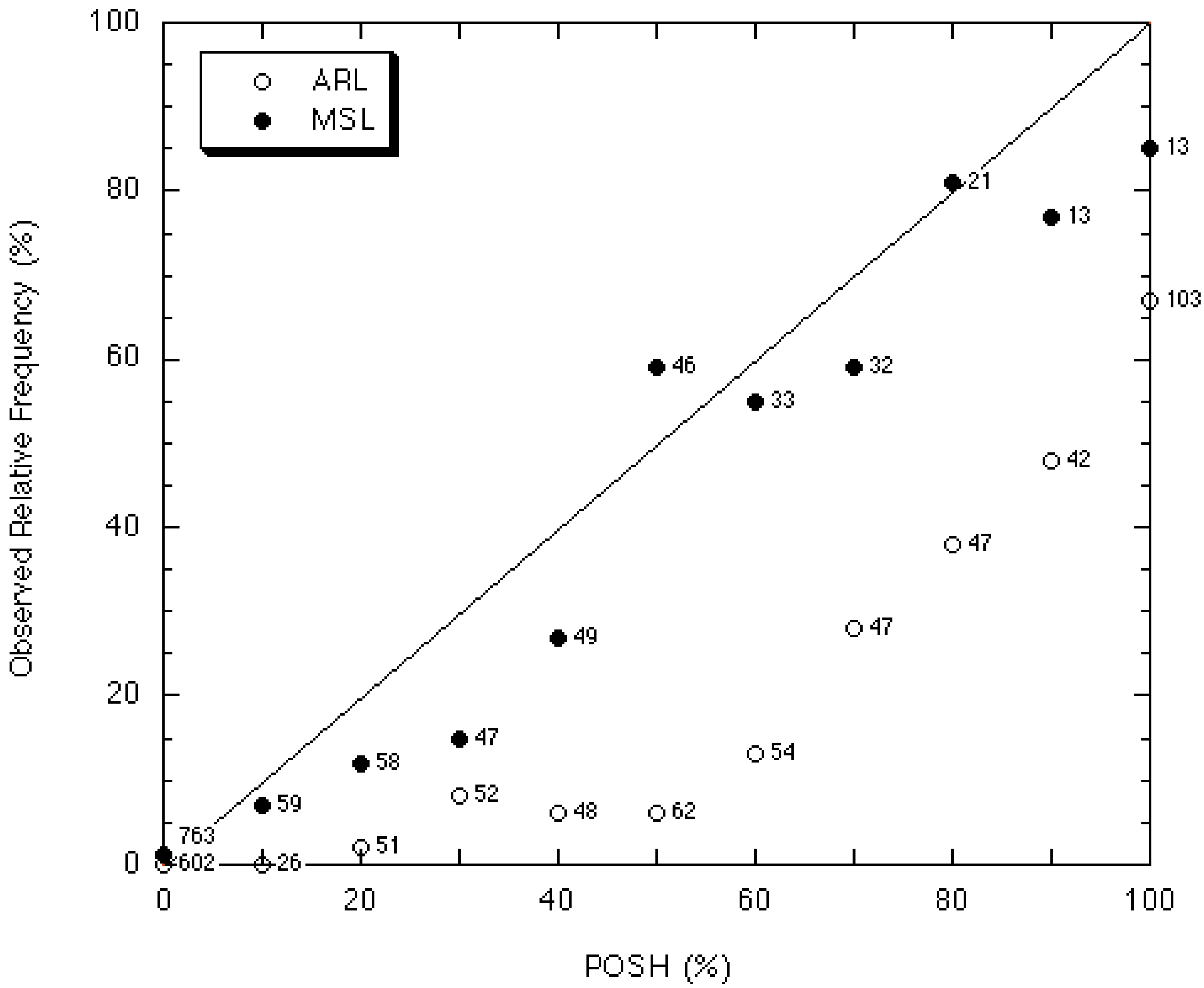

Figure 2

Two days from Montana